

# CPRE

CONSORTIUM *for* POLICY RESEARCH *in* EDUCATION  
UNIVERSITY OF PENNSYLVANIA

UNIVERSITY OF  
DELAWARE

Center for Research in  
Education & Social Policy

# READING RECOVERY: An Evaluation of the Four-Year i3 Scale-Up

A RESEARCH REPORT

Henry May

Philip Sirinides

Abigail Gray

Heather Goldsworthy

MARCH 2016



© Consortium for Policy Research in Education, University of Pennsylvania

The research reported here was supported by the Office of Innovation and Improvement (OII), U.S. Department of Education, through Grant #U396A100027 to The Ohio State University, and in part by the Institute of Education Sciences (IES), U.S. Department of Education, through Grant #R305B090015 to the University of Pennsylvania. The opinions expressed are those of the authors and do not represent the views of OII, IES, or the U.S. Department of Education.

This CPRE research report was externally and internally reviewed for quality and accuracy.

Suggested Citation:

May, H., Sirinides, P., Gray, A., & Goldsworthy, H. (2016). *Reading Recovery: An Evaluation of the Four-Year i3 Scale-Up*. Philadelphia: Consortium for Policy Research in Education.

# Executive Summary

This report presents the final results of a four-year independent external evaluation of the impacts and implementation of the scale-up of Reading Recovery, a literacy intervention targeting struggling 1st-grade students. The evaluation was conducted by the Consortium for Policy Research in Education (CPRE) at the University of Pennsylvania in collaboration with the Center for Research in Education and Social Policy (CRESP) at the University of Delaware.

The evaluation was funded by an Investing in Innovation (i3) grant to The Ohio State University (OSU) from the U.S. Department of Education's Office of Innovation and Improvement. The grant, which was awarded in 2010, totaled \$45 million, with an additional \$10.1 million raised from private sources, to cover the expansion of Reading Recovery around the U.S. Of this, \$5 million was earmarked for the completion of the independent evaluation of the scale-up effort between 2011 and 2015.

Reading Recovery is an intensive intervention targeting struggling 1<sup>st</sup>-grade readers. The program was developed in the 1970s and 80s by Marie Clay, a developmental psychologist and professor at the University of Auckland whose theories about how children become literate provide the foundation for Reading Recovery's approach (Clay, 1991; 2005). Reading Recovery is delivered to students through a 12- to 20-week series of individual, daily, 30-minute lessons provided by a Reading Recovery-trained teacher. The overarching goal of the program is to intervene early, before young students' reading difficulties become lifelong obstacles.

## STUDY GOALS AND DESIGN

CPRE/CRESP's evaluation of Reading Recovery includes parallel rigorous experimental and quasi-experimental designs for estimating program impacts, coupled with a large-scale, mixed-methods study of program implementation under the i3 scale-up. The primary goals of the evaluation are to:

1. Provide experimental evidence of the short- and long-term impacts of Reading Recovery on student learning in schools that are part of the i3 scale-up; and
2. Assess the implementation of Reading Recovery under the i3 grant, including fidelity to the program model and progress toward the scale-up goals.

## KEY FINDINGS

The impact evaluation includes a multi-site randomized controlled trial (RCT) for estimating immediate impacts, a regression discontinuity study (RD) for estimating long-term impacts, and an implementation study for assessing fidelity of implementation and exploring program implementation in depth. The RCT includes nearly 7,000 randomized students in more than 1,200 schools over four years. The RD study measures Reading Recovery's impacts at the end of first grade and in third grade, and replicates the RCT's immediate post-treatment findings in a separate sample of students. The implementation study involves a combination of qualitative and quantitative research executed on a large scale over the same four-year timeframe.

## KEY FINDINGS

The evaluation's key findings pertain to the following topics:

### **1. Scale-Up Processes, Challenges, and Outcomes**

CPRE's evaluation of the i3 scale-up assessed Reading Recovery's progress toward the goals of training 3,675 new Reading Recovery teachers; providing one-to-one Reading Recovery lessons to an additional 67,264 students; and, delivering other instruction—generally classroom or small-group instruction—to 302,688 more students via teachers trained in Reading Recovery during the scale-up. We found that the scale-up surpassed its goals for the number of Reading Recovery teachers trained (3,747, or 102 percent of the scale-up goal), and the number of non-Reading Recovery students served by those teachers (325,458, or 108 percent of the scale-up goal). In the third area, students served with one-to-one Reading Recovery lessons provided by teachers trained with i3 funds, the project achieved 92 percent of its goal, with 61,992 students. As these results indicate, the effort to expand Reading Recovery under the i3 initiative was highly successful.

Our research into the scale-up process revealed that Reading Recovery trainers at the 19 universities that partnered with OSU to implement the scale-up nationally played a critical role in recruiting teachers and schools to participate, as did Reading Recovery teacher leaders across the country. The particular strategies trainers and teacher leaders used in recruitment were varied and context-driven, though the major obstacles—limited staff capacity, a difficult fiscal climate, and the challenges of scaling up sustainably and with fidelity to the Reading Recovery model—were consistent across most regions.

### **2. Immediate Impacts of Reading Recovery**

The four-year, multi-site RCT examined Reading Recovery's impacts on students' scores on the Iowa Test of Basic Skills (ITBS) Reading Total assessment as well as the ITBS Reading Comprehension and Reading Words subtests, and on the Observation Survey of Early Literacy Assessment (OS). Within schools, students were matched into pairs based on pretest scores and randomly assigned, within matched pairs, to treatment and control. Students in the treatment group received 12- to 20-weeks of daily, one-to-one Reading Recovery lessons provided by a trained teacher as a supplement to regular classroom

literacy instruction. Students in the control group received regular classroom literacy instruction as well as any interventions normally provided to low-performing 1st-grade readers in their schools.

The RCT revealed medium to large impacts across all outcome measures. Effect sizes on the ITBS and its subscales at the end of 12- to 20-weeks of treatment ranged between 0.30 and 0.42 standard deviations. For the ITBS Total Reading battery, this effect size translates to a gain of +18 percentage points in the treatment group, as compared with control students. The growth rate we observed in students who participated in Reading Recovery over approximately a five-month period was 131 percent of the national average rate for 1st-grade students. Moreover, these results were similar in two subgroups of interest to the i3 program: English Language Learners and students in rural schools.

With 6,888 student participants, the RCT of short-term impacts in the scale-up schools is among the largest such studies ever conducted. Its rigorous design and large sample offer strong evidence of the effects of Reading Recovery on the short-term progress of struggling students.

### 3. Sustained Impacts of Reading Recovery

The evaluation's RD study focused on students who were in first grade in 2011-2012 and had reached third grade by the end of the i3 study. It assessed these students' reading achievement at the end of first grade, roughly 5 months after they completed Reading Recovery, and again in third grade. Findings at the end of first grade revealed an effect of +12 points on students' OS scores, for an intent to treat effect size of .30 standard deviations.

The RD analysis of impacts on 3<sup>rd</sup>-grade reading achievement used state test scores in reading as the outcome measure. While the impact estimate produced by this analysis was not significant, the available data were far too sparse to produce a conclusive finding. Future research is planned to better address Reading Recovery's impacts on 3<sup>rd</sup>-grade test score.

The RD study also replicated the findings of the RCT. Its baseline contrast was estimated simultaneously with the Year One RCT in a separate random sample of i3 schools, and it found similarly large impacts. Based on a sample of 3,522 students in 274 schools, the impacts estimated immediately post intervention show an effect of +29.7 points on the OS. This equates to an intent to treat effect size of .68 standard deviations.

### 4. Implementation Fidelity

CPRE/CRESP assessed fidelity to the Reading Recovery program model in the schools involved in the scale-up over the four years of the evaluation. Fidelity data were collected annually from Reading Recovery teachers and teacher leaders, and analyzed for consistency with the *Standards and Guidelines of Reading Recovery in the United States, 6<sup>th</sup> Edition* (Reading Recovery Council of North America, 2012). Specifically, we assessed



## CONCLUSIONS

fidelity to the program model in four areas: staff background and selection; teacher leader and site capacity; Reading Recovery teacher training and professional development; and one-to-one Reading Recovery lessons.

Our analysis revealed strong fidelity to the program model in all of these areas and all years of the scale-up. This suggests that the intervention was delivered as designed to the students in the scale-up, and that teachers delivering Reading Recovery lessons were properly trained. In total, the results of the fidelity analysis support the validity of our impact findings.

### 5. Lesson- and School-Level Implementation

We also studied aspects of Reading Recovery's implementation that are not codified in program documents but that vary from school to school and impact program quality. These aspects are manifested both at the level of the one-to-one Reading Recovery lesson and at the level of the school as a whole. First, at the lesson level, we explore instructional strength in terms of Reading Recovery teachers' *deliberateness* and *instructional dexterity*. These related but distinct dimensions of Reading Recovery teachers' instruction are enacted through specific aspects of their practice. We also discuss school- and district-level administrative support as important facilitators of instructional strength in Reading Recovery.

Second, we present the results of 23 field-based case studies focused on schools implementing Reading Recovery under the scale-up, and identify four schemas of Reading Recovery implementation at the school level: isolation, obstruction, endorsement, and integration. Patterns of communication about Reading Recovery within the school and the principal's engagement with the program are the key differentiators of these four schemas.

## CONCLUSIONS

The i3-funded scale-up of Reading Recovery was one of the most ambitious and well-documented expansions of an instructional program in U.S. history, and it was highly successful. The independent evaluation investigating its impacts and execution is one of the most comprehensive evaluations ever implemented in the field of education. It represents a contribution to the growing body of research on the conditions for and impacts of scaling up instructional programs.



